



Matrizes

ISSN: 1982-2073

ISSN: 1982-8160

matrizes@usp.br

Universidade de São Paulo

Brasil

LÉVY, PIERRE

IEML: rumo a uma mudança de paradigma na Inteligência Artificial a

Matrizes, vol. 16, núm. 1, 2022, Enero-, pp. 11-34

Universidade de São Paulo

São Paulo, Brasil

DOI: <https://doi.org/10.11606/issn.1982-8160.v16i1p11-34>

Disponível em: <https://www.redalyc.org/articulo.oa?id=143071289015>

- [Cómo citar el artículo](#)
- [Número completo](#)
- [Más información del artículo](#)
- [Página de la revista en redalyc.org](#)

redalyc.org

Sistema de Información Científica Redalyc

Red de Revistas Científicas de América Latina y el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

IEML: Rumo a uma Mudança de Paradigma na Inteligência Artificial^a

IEML: Towards a Paradigm Shift in Artificial Intelligence

PIERRE LÉVY^b

Université de Montréal, Faculté des Arts et des Sciences, Département de Littératures et de Langues du Monde. Montreal – PQ, Canadá

RESUMO

O objetivo deste ensaio é apresentar uma visão geral das limitações da Inteligência Artificial (IA) contemporânea e propor uma abordagem para superá-las com uma metalinguagem semântica computável. Proponho que a IA adote um modelo computável e univocal da linguagem humana, a MetaLinguagem da Economia da Informação, um código semântico de minha própria invenção que tem o poder expressivo de uma linguagem natural e a sintaxe de uma linguagem regular. Isso pode abrir novos caminhos para a IA criar uma sinergia entre a democratização do controle de dados e o aprimoramento da inteligência coletiva.

Palavras-chave: Inteligência Artificial, código semântico, inteligência coletiva, MetaLinguagem da Economia da Informação

^aTrabalho originalmente publicado em INTLEKT Metadata (2022, 18 de janeiro). <https://bit.ly/3xDMd9V>.

^bProfessor associado na Université de Montréal (Canadá), membro da Royal Society of Canada. Autor de Inteligência Coletiva (1994), Cibercultura (1997), A Esfera Semântica (2011), entre outros livros traduzidos em várias línguas. Orcid: <https://orcid.org/0000-0003-1599-221X>. E-mail: pierre.levy@umontreal.ca

ABSTRACT

The goal of this essay is to present an overview of the limitations of contemporary AI (artificial intelligence) and to propose an approach to overcome them with a computable semantic metalanguage. I propose that AI adopts a computable and univocal model of the human language, the Information Economy Metalanguage (IEML), a semantic code of my own invention. IEML has the expressive power of a natural language and the syntax of a regular language. This can open new avenues for Artificial Intelligence to create a synergy between the democratization of data control and the enhancement of collective intelligence.

Keywords: Artificial Intelligence, semantic code, collective intelligence, Information Economy MetaLanguage

VAMOS PRIMEIRO EXAMINAR como o termo *inteligência artificial* (IA) é usado na sociedade em geral, por exemplo, no jornalismo e na publicidade. A observação histórica mostra a tendência de classificar as aplicações *avançadas* em inteligência artificial nas épocas em que emergem pela primeira vez; no entanto, anos depois, essas mesmas aplicações são frequentemente atribuídas à computação cotidiana. Por exemplo, o reconhecimento de caracteres visuais, originalmente conhecido como sendo IA, agora é considerado comum e muitas vezes é integrado em programas de software sem alarde. Uma máquina capaz de jogar xadrez foi celebrada como uma conquista técnica na década de 1970, mas hoje é possível baixar facilmente um programa de xadrez gratuito no smartphone sem nenhum toque de espanto. Além disso, a depender da IA estar na moda (como hoje) ou desacreditada (como nos anos 1990 e 2000), as estratégias de marketing enfatizarão o termo IA ou o substituirão por outros. Por exemplo, os *sistemas especializados* da década de 1980 tornam-se as inócuas *regras de negócios* nos anos 2000. É assim que técnicas ou conceitos idênticos mudam de nome, conforme a moda, tornando a percepção de seu âmbito e evolução particularmente opaca

Vamos agora deixar o vocabulário do jornalismo ou do marketing para investigar a disciplina acadêmica. Desde a década de 1950, o ramo específico da ciência da computação que se preocupa com a modelagem e simulação de inteligência humana é chamado de Inteligência Artificial.

A modelagem computacional da inteligência humana é um objetivo científico digno, que teve, e continuará a ter, consideráveis benefícios teóricos e práticos. No entanto, a maioria dos pesquisadores da área não acredita que máquinas inteligentes *autônomas* serão construídas em breve, a despeito das previsões iniciais e entusiasmadas sobre a capacidade da IA declaradas nos primeiros anos, contrariadas, mais tarde, pelos fatos. Grande parte da pesquisa nesse campo – e a maioria de suas aplicações práticas – visa aumentar a *cognição humana* em vez de reproduzi-la mecanicamente. Isso contrasta com o programa de pesquisa focado na construção de uma Inteligência Artificial geral autônoma.

Defendi a ideia de Inteligência Artificial a serviço da inteligência coletiva e do desenvolvimento humano no livro *La Sphère Sémantique* (Lévy, 2011). Vamos continuar essa linha de pensamento neste ensaio.

Do ponto de vista técnico, a IA é dividida em dois ramos principais: estatístico e simbólico. Um algoritmo estatístico de IA *aprende* com os dados fornecidos. Simula, portanto, (imperfeitamente, como veremos na sequência) a dimensão *indutiva* do raciocínio humano. Em contraste, a IA simbólica não aprende com os dados, mas depende da formalização lógica de um domínio do conhecimento como projetado pelos engenheiros. Em princípio, comparada com a IA estatística, exige,

portanto, uma maior quantidade de trabalho intelectual humano. Um algoritmo simbólico de IA aplica as regras que recebeu aos dados fornecidos. Assim, simula mais da dimensão *dedutiva* do raciocínio humano. Revisarei sucessivamente esses dois ramos principais da IA, com um foco especial em destacar suas limitações.

IA E SUAS LIMITAÇÕES

IA Neural

O ramo *estatístico* da IA envolve o treinamento de algoritmos a partir de acumulações massivas de dados para permitir o reconhecimento visual, de áudio, linguístico e de outras formas de informação. Isso se chama *aprendizado de máquina*. Quando falamos de IA em 2022, geralmente designamos esse tipo de programa de pesquisa técnica e científica. Como observamos, a IA estatística utiliza o trabalho humano com moderação em comparação com a IA simbólica. Em vez de ter que escrever um programa de reconhecimento de padrões, basta fornecer um conjunto de dados de treinamento para o algoritmo de aprendizagem de máquina. Se, por exemplo, a IA estatística receber milhões de imagens de patos com rótulos especificando que a imagem representa um pato, ela aprende a reconhecer um pato e, após a conclusão de seu treinamento, será capaz de afixar o rótulo *pato* em uma imagem não categorizada dessa ave. Ninguém explicou à máquina *como* reconhecer um pato: basta apenas dar exemplos. A tradução automática funciona com o mesmo princípio: uma IA estatística recebe milhões de textos na linguagem A, acompanhados de sua tradução para a linguagem B. Treinado com esses exemplos, o sistema aprende a traduzir um texto da linguagem A para a linguagem B. É assim que algoritmos de tradução automática como DeepL ou Google Translate funcionam. Para dar um exemplo de outro campo, a IA estatística usada para dirigir *veículos autônomos* também funciona *combinando dois conjuntos de dados*: as imagens da estrada são combinadas com *ações* como aceleração, frenagem, giro etc. Em suma, a IA estatística estabelece uma conexão (*mapeamento*) entre um conjunto de dados e um conjunto de rótulos (no caso de reconhecimento de padrões) ou entre dois conjuntos de dados (no caso da tradução ou dos veículos autônomos). A IA estatística, portanto, destaca-se na categorização, reconhecimento de padrões e correspondência entre dados perceptivos e motores.

Em sua versão mais avançada, a IA estatística é baseada em modelos de *rede neural* que simulam aproximadamente a maneira como o cérebro aprende. Esses modelos são chamados de *aprendizagem profunda* porque são baseados

na sobreposição de *múltiplas camadas* de neurônios formais. As redes neurais são o subcampo mais complexo e avançado da IA estatística. Esse tipo neural de inteligência artificial existe desde a origem da ciência da computação, como ilustrado pela pesquisa de McCulloch nas décadas de 1940 e 1950, Frank Rosenblatt e Marvin Minsky nos anos 1950, e von Foerster nas décadas de 1960 e 1970. Trabalhos significativos nessa área também foram feitos na década de 1980, especialmente envolvendo David Rumelhart e Geoffrey Hinton, mas toda essa pesquisa teve pouco sucesso prático até a década de 2010.

Além de certos refinamentos científicos dos modelos, dois fatores independentes dos avanços teóricos explicam o uso crescente de redes neurais: a disponibilidade de enormes quantidades de dados e o aumento do poder computacional. A partir da segunda década do século XXI, as organizações estão engajadas na transformação digital, e uma parcela crescente da população mundial está usando a web. Tudo isso gera fluxos de dados gigantescos. As informações produzidas, portanto, são processadas por grandes plataformas digitais em *data centers* (a *nuvem*) que concentram um poder computacional sem precedentes. No início do século XXI, as redes neurais foram implementadas por processadores originalmente projetados para a computação gráfica, mas, hoje em dia, os *data centers* de propriedade das grandes empresas de tecnologia já usam processadores especificamente projetados para aprendizado neural. Assim, modelos teóricos interessantes, mas impraticáveis, do século XX, tornaram-se subitamente bastante relevantes no século XXI, a ponto de apoiar uma nova indústria.

Diminuindo os Retornos

No entanto, após o turbilhão de avanços no aprendizado de máquina baseado em redes neurais na década de 2010, o progresso parece ter parado nos últimos anos. De fato, para obter agora um desempenho mesmo marginalmente aprimorado, o tamanho dos conjuntos de dados e o poder computacional usado para treinar modelos devem ser multiplicados por várias ordens de magnitude. Já chegamos à era do *decréscimo dos retornos cognitivos* da IA neural; portanto, chegou a hora de questionar as limitações desse conjunto de técnicas e vislumbrar seriamente uma mudança de paradigma.

Os principais problemas repousam na qualidade dos dados de treinamento, na falta de modelagem causal, na natureza inexplicável de alguns dos resultados, na ausência de generalização, no significado supostamente inescrutável dos dados e, finalmente, nas dificuldades em acumular e integrar o conhecimento.

Qualidade dos Dados de Treinamento

Um engenheiro do Google é citado como se dissesse, brincando: “Toda vez que despedimos um linguista, nosso desempenho de tradução automática melhora”. Mas, embora a IA estatística seja conhecida por ter pouca necessidade de trabalho humano, os riscos de viés e erro apontados por usuários cada vez mais preocupados estão impulsionando a necessidade de uma melhor seleção de dados de treinamento, incluindo uma rotulagem mais cuidadosa. No entanto, isso requer tempo e experiência humana, precisamente os fatores que se espera eliminar.

Ausência de uma Hipótese Causal Explícita

Todos os cursos de estatística começam com um alerta sobre a confusão entre correlação e causalidade. Uma correlação entre A e B não prova que A é a causa de B. Pode ser uma coincidência, ou B pode ser a causa de A, ou mesmo um fator C não considerado pela coleta de dados é a verdadeira causa de A e B, sem mencionar todas as complexas relações sistêmicas imagináveis envolvendo A e B. No entanto, o aprendizado de máquina é baseado na correspondência de conjuntos de dados por meio de correlações. A noção de causalidade é estranha à IA estatística, como acontece com muitas técnicas usadas para analisar coletas massivas de dados, embora as suposições causais estejam muitas vezes implícitas na escolha dos conjuntos de dados e sua categorização. Em resumo, a IA neural/estatística contemporânea não é capaz de distinguir causa e efeito. Até agora, ao usar a IA para auxiliar na tomada de decisões e, de modo mais geral, para orientação em domínios práticos, modelos causais explícitos são indispensáveis, pois, para que as ações sejam eficazes, elas devem intervir *nas causas*.

Em uma abordagem científica integral, as medições estatísticas e as hipóteses causais trabalham em controle uníssono e recíproco. Mas considerar apenas correlações estatísticas criaria um perigoso ponto cego cognitivo. Já a prática generalizada de manter as teorias causais implícitas impede relativizá-las, compará-las com outras teorias, generalizá-las, compartilhá-las, criticá-las e melhorá-las.

Resultados Inexplicáveis

O funcionamento das redes neurais é opaco. Milhões de operações transformam gradualmente a força das conexões dos conjuntos neurais que são feitos de centenas de camadas.

Uma vez que os resultados dessas operações não podem ser explicados ou justificados conceitualmente de uma forma que os seres humanos possam

entender, é difícil confiar nesses modelos. Essa falta de explicação torna-se preocupante quando as máquinas tomam decisões financeiras, legais, médicas ou conduzem veículos de forma autônoma, sem mencionar aplicações militares. Para superar esse obstáculo, e paralelamente ao desenvolvimento de uma inteligência artificial mais ética, cada vez mais pesquisadores estão explorando o novo campo de pesquisa da *IA explicativa*.

A Falta de Generalização

À primeira vista, a IA estatística apresenta-se como uma forma de raciocínio indutivo, ou seja, como uma capacidade de inferir regras gerais de uma infinidade de casos. No entanto, os sistemas contemporâneos de aprendizagem de máquina não conseguem generalizar além dos limites dos dados de treinamento que lhes foram fornecidos. Nós – humanos – não só somos capazes de generalizar a partir de alguns exemplos, enquanto é preciso milhões de casos para treinar máquinas, mas podemos abstrair e conceituar o que aprendemos, enquanto o aprendizado de máquina não consegue extrapolar, muito menos, conceituar. A IA estatística permanece no nível de aprendizagem puramente reflexa, sua generalização é estritamente circunscrita aos exemplos fornecidos com os quais é alimentada.

Significado Inacessível

Enquanto o desempenho na tradução e na escrita automáticas (como ilustrado pelo programa GPT3) está avançando, as máquinas ainda não entendem o *significado* dos textos que traduzem ou escrevem. Suas redes neurais se assemelham ao cérebro de um papagaio mecânico, capaz apenas de imitar o desempenho linguístico sem entender nada do conteúdo dos textos que está traduzindo. Em poucas palavras, a Inteligência Artificial contemporânea pode aprender a traduzir textos, mas é incapaz de aprender qualquer coisa com essas traduções.

O Problema da Acumulação e da Integração do Conhecimento em IA Estatística

Desprovida de conceitos, a IA estatística tem dificuldade em *acumular* conhecimento. Por isso, a *integração* do conhecimento de diversas áreas da expertise parece fora de alcance. Essa situação não favorece a troca de conhecimento entre máquinas. Portanto, muitas vezes é necessário começar do zero para cada novo projeto. No entanto, devemos apontar a existência de modelos de processamento de linguagem natural, como o BERT, que são pré-treinados

em dados gerais e têm então a possibilidade de se especializar. Uma forma de capitalização é possível, de maneira limitada. Mas continua impossível integrar todo o conhecimento objetivo acumulado ao longo dos séculos pela humanidade em um sistema neuromimético.

IA SIMBÓLICA E SEUS LIMITES

Durante os últimos setenta anos, o ramo *simbólico* da IA tem correspondido sucessivamente ao que é conhecido como: redes semânticas, sistemas baseados em regras, bases de conhecimento, sistemas especializados, web semântica e, mais recentemente, gráficos de conhecimento. Desde suas origens, nos anos 1940-50, boa parte da ciência da computação de fato pertence à IA simbólica.

A IA simbólica codifica explicitamente o conhecimento humano na forma de redes de relações entre categorias e regras lógicas que permitem o raciocínio automático. Seus resultados são, portanto, mais facilmente explicados do que os da IA estatística.

A IA simbólica funciona bem nos micromundos fechados de jogos ou laboratórios, mas rapidamente se torna sobrecarregada em ambientes abertos que não seguem um número limitado de regras rígidas. A maioria dos programas simbólicos de IA usados em ambientes de trabalho do mundo real resolvem problemas apenas em um domínio bastante restrito, no diagnóstico médico, na resolução de problemas de máquina, conselhos de investimento etc. Um *sistema especializado*, na verdade, funciona como um meio para o encapsulamento e distribuição de determinado know-how que pode ser utilizado onde for necessário. A habilidade prática então se torna disponível mesmo na ausência de conhecimento humano.

No final da década de 1980, após uma série de promessas muito otimistas seguidas de decepções, começou o que tem sido chamado de *inverno* da inteligência artificial (combinando todas as tendências). No entanto, os mesmos processos continuam a ser aplicados para resolver os mesmos tipos de problemas indiscriminadamente, só abandonamos o programa de pesquisa geral no qual esses métodos foram incorporados. Assim, no início do século XXI, as regras comerciais do software corporativo e as ontologias da *Web Semântica* sucederam os sistemas especializados da década de 1980. Apesar das mudanças de nome, é fácil reconhecer nessas novas especialidades os antigos processos de IA simbólica.

No início dos anos 2000, a Web Semântica tem como objetivo explorar todas as informações disponíveis na Web. Para tornar os dados legíveis pelos computadores, diferentes domínios de conhecimento ou prática são organizados em modelos coerentes. Essas são as *ontologias*, que só podem reproduzir a

compartimentação lógica das décadas anteriores, mesmo que os computadores estejam agora muito mais interligados.

Infelizmente, encontramos na IA simbólica as mesmas dificuldades na integração e acúmulo de conhecimento que na IA estatística. Essa compartimentação está em oposição ao projeto original da Inteligência Artificial como disciplina científica, que quer modelar a inteligência humana em geral, e que normalmente tende a uma acumulação e integração de conhecimentos que podem ser mobilizados pelas máquinas.

Apesar da compartimentação de seus modelos, a IA simbólica é, no entanto, ligeiramente melhor do que a IA estatística em termos de acumulação e troca de dados. Um número crescente de empresas, começando pelas Big Tech, estão organizando suas bases de dados usando um gráfico de conhecimento que está constantemente sendo melhorado e aumentado.

Além disso, o Wikidata oferece um bom exemplo de um gráfico de conhecimento aberto, por meio do qual as informações que vão gradualmente se acumulando podem ser lidas tão bem pelas máquinas quanto pelos humanos. No entanto, cada um desses gráficos de conhecimento é organizado de acordo com os propósitos – sempre particulares – de seus autores e não pode ser facilmente reutilizado para outros fins. Nem a IA estatística nem a IA simbólica possuem as propriedades da recombinação maleável que devemos esperar justamente dos módulos de uma Inteligência Artificial a serviço da inteligência coletiva.

A IA Simbólica é uma Consumidora Voraz do Trabalho Intelectual Humano

Houve muitas tentativas de conter todo o conhecimento humano em uma única ontologia para permitir uma melhor interoperabilidade, mas então a vibração, complexidade, evolução e múltiplas perspectivas do conhecimento humano são apagadas. Em um nível prático, ontologias universais – ou mesmo aquelas que afirmam formalizar todas as categorias, relações e regras lógicas de um vasto domínio – rapidamente se tornam enormes, complicadas e difíceis de entender e manter para o ser humano que deve lidar com elas. Um dos principais gargalos da IA simbólica é a quantidade e a alta qualidade do trabalho humano necessários para modelar um domínio do conhecimento, ainda que bastante circunscrito. De fato, é necessário não só ler a literatura, mas também entrevistar e ouvir longamente vários especialistas do domínio a ser modelado. Adquirido pela experiência, o conhecimento desses especialistas é mais frequentemente expresso por meio de histórias, exemplos e descrições de situações típicas. É então necessário transformar o conhecimento empírico e oral em um modelo lógico coerente cujas regras devem ser executáveis por um computador. Eventualmente, o raciocínio dos

especialistas será automatizado, mas o trabalho de *engenharia do conhecimento* a partir do qual a modelagem prossegue não pode ser.

POSIÇÃO PROBLEMÁTICA: QUAL É O PRINCIPAL OBSTÁCULO PARA (MAIS) DESENVOLVIMENTO DA IA?

Rumo a uma Inteligência Artificial Neuro-Simbólica

Agora é hora de dar um passo atrás. Os dois ramos da IA – neural e simbólico – existem desde meados do século XX e correspondem a dois estilos cognitivos igualmente presentes em humanos. Por um lado, temos o reconhecimento de padrões, que corresponde aos módulos sensório-motores reflexos, sejam eles aprendidos ou de origem genética. Por outro lado, temos o conhecimento conceitual explícito e reflexivo, muitas vezes organizado em modelos causais e que podem ser objeto de raciocínio.

Uma vez que esses dois estilos cognitivos trabalham juntos na cognição humana, não há razão teórica para não tentar fazê-los cooperar em sistemas de Inteligência Artificial. Os benefícios são óbvios e cada um dos dois subsistemas pode sanar problemas encontrados pelo outro. Em uma IA mista, o componente simbólico supera as dificuldades de conceituação, generalização, modelagem causal e transparência do componente neural. Simetricamente, o componente neural traz as capacidades de reconhecimento de padrões e aprendizado a partir de exemplos que estão faltando na IA simbólica.

Importantes pesquisadores de IA, bem como muitos observadores reconhecidos da disciplina, estão gravitando na direção de uma IA híbrida. Por exemplo, Dieter Ernst (2021) defendeu recentemente uma “integração entre redes neurais, que se destacam na classificação perceptiva e sistemas simbólicos, que por sua vez se destacam na abstração e inferência” (“AI Research and Governance Are at a Crossroads”, para. 4).

Seguindo os passos de Gary Marcus, os pesquisadores da IA Luis Lamb e Arthur d’Avila Garcez publicaram recentemente um artigo em favor de uma IA neuro-simbólica na qual representações adquiridas por meios neurais seriam interpretadas e processadas por meios simbólicos. Parece que encontramos uma solução para o problema de interrupção no desenvolvimento da IA: seria benéfico acoplar, de modo inteligente, os ramos simbólico e estatístico, em vez de mantê-los separados como dois programas de pesquisa concorrentes. Além disso, não vemos as empresas de Big Tech, que destacam o aprendizado de máquina e a IA neural em seus esforços de relações públicas, desenvolvendo internamente, de

forma discreta, gráficos de conhecimento para organizar sua memória digital e dar sentido aos resultados de suas redes neurais? Mas antes de declararmos a questão resolvida, vamos pensar um pouco mais sobre os dados do problema.

Cognição Animal e Cognição Humana

Para cada um dos dois ramos da IA, listamos os obstáculos que impedem uma Inteligência Artificial menos fragmentada, mais útil e mais transparente. No entanto, encontramos a mesma desvantagem de ambos os lados: a compartimentação lógica e as dificuldades de acumulação e integração. Combinar o neural e o simbólico não nos ajudará a superar esse obstáculo, já que nenhum deles pode fazê-lo. No entanto, as sociedades humanas reais podem transformar percepções tácitas e habilidades experienciais em conhecimento compartilhável. Por meio de um amplo diálogo, um especialista em uma área acabará por se fazer entender por um especialista em outro campo e pode até mesmo lhe ensinar algo. Como esse tipo de desempenho cognitivo pode ser reproduzido em sociedades de máquinas? Qual fator desempenha o papel integrativo da linguagem natural nos sistemas de Inteligência Artificial?

Muitas pessoas pensam que, como o cérebro é o recipiente orgânico da inteligência, modelos neurais são a chave para sua simulação. Mas de que tipo de inteligência estamos falando? Não vamos esquecer que todos os animais têm um cérebro, e não é a inteligência do mosquito ou da baleia que a IA quer simular, mas a do ser humano. E se somos *mais inteligentes* do que outros animais (pelo menos do nosso ponto de vista) não é por causa do tamanho do nosso cérebro. Os elefantes têm cérebros maiores que humanos em termos absolutos, e a proporção do tamanho do cérebro para o tamanho do corpo é maior em camundongos do que em humanos. É principalmente nossa capacidade linguística, predominantemente processada nas áreas de Broca e Wernicke do cérebro (exclusivas da espécie humana), que distingue nossa inteligência da de outros vertebrados superiores. No entanto, esses módulos de processamento de linguagem não são funcionalmente separados do resto do cérebro; ao contrário, informam todos os nossos processos cognitivos, incluindo nossas habilidades técnicas e sociais. Nossas percepções, ações, emoções e comunicações são linguisticamente codificadas, e nossa memória é em grande parte organizada por um sistema de semântica coordenada fornecido pela linguagem.

Tudo bem, pode-se dizer. Simular habilidades de processamento simbólico humano, incluindo a faculdade linguística, não é precisamente o que a IA simbólica deve fazer? Mas então por que a IA está compartimentalizada em ontologias distintas, ainda que lutando para garantir a interoperabilidade semântica de seus

sistemas, e tem muita dificuldade em acumular e intercambiar conhecimento? Simplesmente porque, apesar de seu nome de *simbólica*, a IA ainda não tem um modelo computável de linguagem. Desde o trabalho de Chomsky, sabemos como calcular a dimensão sintática das línguas, mas sua dimensão semântica permanece fora do alcance da ciência da computação. Para entender essa situação, é necessário recordar alguns elementos da semântica.

Semântica em Linguística

Do ponto de vista do estudo científico da linguagem, a semântica de uma palavra ou uma frase pode ser dividida em duas partes que são combinadas na prática, embora conceitualmente distintas: semântica linguística e semântica referencial. A semântica linguística lida com a relação entre palavras, enquanto a referencial está preocupada com a relação entre palavras e coisas.

Semântica Linguística (Palavra-Palavra)

Um símbolo linguístico (palavra ou frase) geralmente tem dois aspectos: o *significante*, que é uma imagem visual ou acústica, e o *significado*, que é um conceito ou uma categoria geral. Por exemplo, o significante *árvore* tem o seguinte significado: “uma planta amadeirada de tamanho variável, cujo tronco cultiva galhos a partir de uma altura específica”. Dado que a relação entre significante e significado é estabelecida por uma língua, o significado de uma palavra ou de uma frase é definido como *um nó de relações com outros significados*. Em um dicionário clássico, cada palavra está situada em relação a outras palavras associadas (o tesouro), e seu significado é explicado por frases (a definição) que usam outras palavras que são explicadas por outras frases, e assim por diante, de forma circular. A *semântica linguística* é fundamental para um dicionário clássico. Verbos e substantivos comuns (por exemplo, *árvore*, *animal*, *órgão*, *comer*) representam categorias conectadas por uma densa rede de *relações semânticas* como: “faz parte”, “é um tipo de”, “pertence ao mesmo contexto que”, “é a causa de”, “é anterior a” etc. Pensamos e nos comunicamos da maneira humana porque nossas memórias coletivas e pessoais estão organizadas em categorias gerais conectadas por relações semânticas.

Semântica Referencial (Palavra-Coisa)

Em contraste com a semântica linguística, a *semântica referencial* faz a ponte entre um *símbolo linguístico* (significante e significado) e um *referencial* (uma instância real). Quando digo que “carvalhos são árvores”, estou especificando o significado convencional da palavra *carvalho* colocando-a em uma

relação espécie-a-gênero com a palavra *árvore*; portanto, estou estritamente colocando a semântica linguística em ação. Mas se eu disser que “aquela árvore no quintal é um carvalho”, então eu estou apontando para uma situação real, e minha proposta é verdadeira ou falsa. Essa segunda declaração, obviamente, coloca a semântica linguística em ação, uma vez que primeiro devo saber o significado das palavras e da gramática para entendê-la. Mas, além da dimensão linguística, a semântica referencial também está envolvida, uma vez que a declaração se refere a um objeto específico em uma situação concreta. Algumas palavras, como substantivos próprios, não têm significado; este refere-se diretamente a um referente. Por exemplo, o significante *Alexandre, o Grande* refere-se a uma figura histórica e o significante *Tóquio*, a uma cidade. Em contraste com um dicionário clássico que define conceitos ou categorias, um dicionário enciclopédico contém descrições de *indivíduos* reais ou fictícios com nomes próprios como divindades, heróis de romances, figuras e eventos históricos, objetos geográficos, monumentos, obras da mente etc. Sua principal função é listar e descrever objetos externos ao sistema de uma língua. Por isso, registra a semântica referencial.

Nota bene: Uma *categoria* é uma classe de indivíduos, uma abstração. Pode haver categorias de entidades, processos, qualidades, quantidades, relações etc. As palavras *categoria* e *conceito* são tratadas aqui como sinônimos.

Semântica em IA

Na ciência da computação, as referências reais, ou indivíduos (as realidades de que falamos) tornam-se os *dados*, enquanto as categorias gerais se tornam os cabeçalhos, campos ou *metadados* usados para classificar e recuperar dados. Por exemplo, no banco de dados de uma empresa, *nome do funcionário*, *endereço* e *salário* são categorias ou metadados, enquanto *Tremblay, 33 Boulevard René Lévesque* e *65 K\$/ano* são dados. Nesse domínio técnico, a semântica referencial corresponde à relação entre dados e metadados e a semântica linguística, à relação entre metadados ou categorias organizadoras, geralmente representadas por palavras ou expressões linguísticas curtas.

Tendo em vista que o objetivo da ciência da computação é aumentar a inteligência humana, uma de suas tarefas deve nos ajudar a entender a inundação de dados digitais e extrair o máximo de conhecimento utilizável possível deles. Para isso, devemos categorizar corretamente os dados – isto é, implementar a semântica da palavra-coisas –, e organizar as categorias de acordo com relações relevantes que nos permitem extrair todo o conhecimento acionável dos dados – o que corresponde à semântica palavra-palavra.

Ao discutir o tema da semântica na ciência da computação, devemos lembrar que os computadores não enxergam espontaneamente uma palavra ou uma frase como um conceito em uma certa relação com outros conceitos em uma língua, mas apenas como uma sequência de letras, ou *sequência de caracteres*. Portanto, as relações entre categorias que parecem óbvias para os seres humanos e que fazem parte da semântica linguística devem ser adicionadas – principalmente à mão – a um banco de dados se um programa deve levá-las em conta.

Vamos agora examinar até que ponto a IA simbólica modela a semântica. Se considerarmos as ontologias da Web Semântica (o padrão em IA simbólica), descobrimos que o significado das palavras não depende da circularidade autoexplicativa da linguagem (como em um dicionário clássico), mas que as palavras apontam para Identificadores de Recursos Uniformes (Uniform Resource Identifiers – URI) na forma da *semântica referencial* (como em um dicionário enciclopédico).

Em vez de confiar em conceitos (ou categorias) que já são dados em uma língua e que aparecem desde o início como nós de relações com outros conceitos, a estrutura da Web Semântica se baseia em conceitos que são definidos separadamente uns dos outros por meio de identificadores únicos. A circulação de sentido em uma rede de significados é descontada em favor de uma relação direta entre significante e referente, como se todas as palavras fossem substantivos próprios. Na ausência de uma semântica linguística baseada em uma gramática e dicionário comuns, as ontologias permanecem assim compartimentalizadas. *Em resumo, a IA simbólica contemporânea não tem acesso ao pleno poder cognitivo e comunicativo da linguagem porque não tem uma linguagem, apenas uma semântica referencial rígida.*

Então, por que a IA não usa línguas naturais – com sua semântica linguística inerente – para representar o conhecimento? A resposta é bem conhecida: as línguas naturais são ambíguas. Uma palavra pode ter vários significados, um significado pode ser expresso por diversas palavras, frases têm múltiplas interpretações possíveis, gramática é elástica etc. Uma vez que os computadores não são seres imbuídos de um bom senso nato, como nós, eles são incapazes de desambiguar corretamente declarações em linguagem natural. Para seus falantes humanos, uma linguagem natural fornece um dicionário, que é uma rede de categorias gerais predefinidas mutuamente explicativas. Essa rede semântica comum permite a descrição e comunicação de múltiplas situações concretas, bem como diferentes domínios do conhecimento. No entanto, devido a suas irregularidades, a IA não pode usar linguagens naturais para se comunicar ou ensinar às máquinas diretamente. É por isso que a IA permanece fragmentada hoje em microdomínios de práticas e conhecimentos, cada um com sua própria semântica particular.

A automação da semântica linguística poderia abrir novos horizontes de comunicação e raciocínio para a Inteligência Artificial. Para lidar com a semântica linguística, a IA precisa de uma linguagem padronizada e univocal, um código especialmente projetado para uso de máquinas e que os humanos poderiam facilmente entender e manipular. Essa linguagem finalmente permitiria que os modelos se conectassem e o conhecimento se acumulasse. Em suma, *o principal obstáculo para o desenvolvimento da IA é a falta de uma linguagem computável comum*. Este é precisamente o problema resolvido pela Metalinguagem da Economia da Informação (Information Economy Metalanguage – IEML), uma metalinguagem que pode expressar significado, como línguas naturais, e cuja semântica é inequívoca e computável, como uma linguagem matemática. O uso do IEML tornará a IA menos onerosa em termos de trabalho humano, mais apta a lidar com significado e causalidade e, o mais importante, capaz de acumular e trocar conhecimentos.

Sem linguagem, não teríamos acesso a questionamento, diálogo ou narrativa. A linguagem é simultaneamente um meio de inteligência pessoal – é difícil pensar sem diálogo interno – e de inteligência coletiva. Grande parte do conhecimento da sociedade foi acumulado e repassado de forma linguística. Dado o papel do discurso na inteligência humana, é surpreendente que esperamos alcançar a inteligência artificial geral sem um modelo computável de linguagem e sua semântica. A boa notícia é que finalmente temos um.

IEML: UMA SOLUÇÃO BASEADA EM UM CÓDIGO SEMÂNTICO

A Metalinguagem da Economia da Informação

Muitos avanços na ciência da computação vêm da invenção de um sistema de codificação relevante que torna o objeto codificado (número, imagem, som etc.) facilmente computável por uma máquina. Por exemplo, codificação binária para números e codificação de pixels ou vetores para imagens. Por isso, temos trabalhado no design de um código que torna a semântica linguística computável. Essa linguagem artificial, IEML, tem uma gramática regular e um dicionário compacto de três mil palavras. Categorias mais complexas podem ser construídas combinando palavras em frases de acordo com um pequeno conjunto de regras gramaticais. Essas categorias complexas, por sua vez, podem ser usadas para definir outras, e assim por diante, recursivamente. Resumindo, qualquer tipo de categoria pode ser construído a partir de um pequeno conjunto de palavras.

Em um nível linguístico, a IEML tem a mesma capacidade expressiva que uma língua natural, e pode ser traduzida em qualquer outra língua. É também uma linguagem univocal: cada palavra do dicionário tem *apenas um significado* (ao contrário das línguas naturais) e um conceito tem apenas uma expressão, tornando sua semântica linguística computável. É importante notar que a IEML não é uma ontologia universal, mas é de fato uma *linguagem* que pode expressar qualquer ontologia ou classificação.

Em um nível matemático, IEML é uma linguagem regular no sentido estabelecido por Chomsky: é uma álgebra. É, portanto, favorável a todos os tipos de processamento automático e transformações.

Em um nível de ciência da computação, como veremos em mais detalhes a seguir, essa metalinguagem fornece uma linguagem de programação especializada para o design de gráficos de conhecimento e modelos de dados.

O Editor IEML

A IEML é definida por sua gramática e dicionário de três mil palavras, que podem ser encontrados no site intlekt.io. Essa metalinguagem vem equipada com uma *ferramenta digital* para facilitar sua escrita, leitura e uso: o editor IEML.

O editor IEML é usado para produzir e explorar *modelos de dados*. Essa noção de *modelo* abrange redes semânticas, sistemas de metadados semânticos, ontologias, gráficos de conhecimento e sistemas de rotulagem para categorizar dados de treinamento. O editor contém uma linguagem de programação para automatizar a criação de nós (categorias) e links (relações semânticas entre categorias). Esta linguagem de programação é *declarativa*, o que significa que não pede ao usuário para organizar um fluxo de instruções condicionais, mas apenas para descrever os resultados desejados.

(1) Com o editor IEML, o modelador humano pode elaborar as *categorias* que servirão como recipientes (ou caixas de memória) para diferentes tipos de dados. Como dito acima, se algumas categorias não podem ser encontradas no dicionário IEML de três mil *palavras*, o modelador pode criar mais delas, combinando palavras para fazer frases, trazendo grande refinamento à categorização.

(2) A partir das categorias, o modelador então programa as *relações semânticas* (*faz parte, é uma causa* etc.) que conectarão os dados categorizados. A ligação entre nós é automatizada com base nas funções gramaticais das categorias. As propriedades matemáticas das relações (reflexividade, simetria, transitividade) são então especificadas.

(3) Uma vez categorizados os dados, o programa tece automaticamente uma rede de relações semânticas, dando aos dados, ao fim, ainda mais significado. A mineração de dados, a exploração hipertextual e a visualização de relacionamentos por tabelas e gráficos permitirão que os usuários finais explorem o conteúdo modelado.

Vantagens

Várias características fundamentais distinguem o editor IEML das ferramentas contemporâneas de modelagem de dados: categorias e relacionamentos são programáveis e os modelos resultantes são interoperáveis e transparentes.

Categorias e relacionamentos são programáveis

A estrutura regular do IEML permite que categorias sejam geradas e as relações sejam tecidas *funcional ou automaticamente*, em vez de serem criadas uma a uma. Essa propriedade economiza tempo considerável para o modelador. O tempo economizado pela automação da criação de categorias e relacionamentos mais do que compensa o tempo gasto na codificação de categorias no IEML, especialmente porque, uma vez criadas, novas categorias e relacionamentos podem ser trocados entre os usuários.

Os modelos são interoperáveis

Todos os modelos são baseados no mesmo dicionário de três mil palavras e estabelecem regras gramaticais. Os modelos são, portanto, *interoperáveis*, o que significa que eles podem facilmente se fundir ou intercambiar categorias e submodelos. Cada modelo ainda é personalizado para um contexto específico, mas os modelos agora podem se comparar, interconectar e integrar.

Os modelos são transparentes

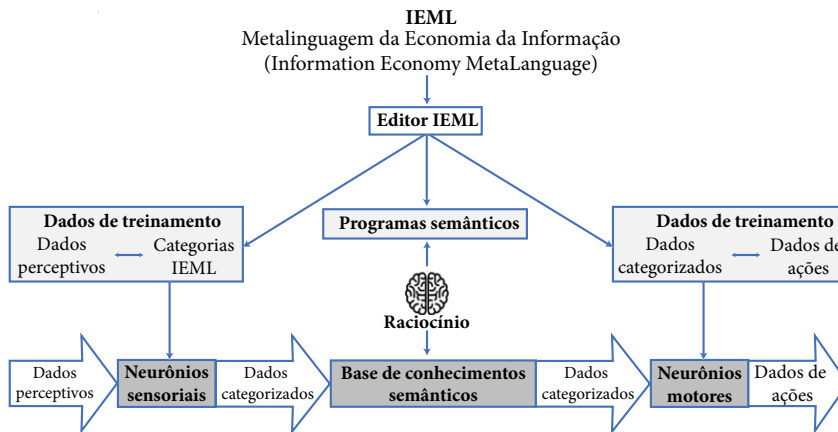
Embora codificados no IEML, os modelos escritos com o editor IEML são *legíveis em linguagem natural*. Uma vez que as categorias e as relações são rotuladas com palavras ou com frases mais elaboradas em linguagens naturais (e sem ambiguidade semântica), os modelos são *mais claros* tanto para os modeladores quanto para os usuários finais, alinhando-se, portanto, aos *princípios contemporâneos de ética e transparência*.

O usuário não precisa ser um cientista da computação ou estar familiarizado com a linguagem IEML para aprender a usá-la com sucesso; a curva de aprendizado é curta. Apenas a gramática (simples e regular) precisa ser dominada. O editor do IEML poderia ser utilizado nas escolas e, portanto, abrindo caminho para uma democratização da alfabetização em dados.

A Arquitetura Neuro-Semântica da IEML

Figura 1

Uma arquitetura neuro-semântica para a IA



Vamos agora propor uma arquitetura (Figura 1) de sistema de IA baseada na IEML, claramente um exemplo particular de uma arquitetura neuro-simbólica, apesar de ser chamada de *neuro-semântica* para enfatizar que resolve o problema da *computação semântica* e da *interoperabilidade semântica* entre sistemas.

Neurônios sensório-motores

Devemos distinguir vários tipos de dados de treinamento (texto, imagem, som etc.), dos quais resultam vários tipos de redes neurais. As redes neurais sensoriais que foram treinadas por meio de exemplos de dados categorizados na IEML inserirão informações no sistema. Os dados categorizados pelos neurônios sensoriais são transmitidos para a base de conhecimentos semânticos. Se forem detectadas inconsistências, erros ou vieses, os dados de treinamento ou sua conceituação devem, obviamente, ser revistos. Assim, o sistema deve incluir um ciclo de diálogo entre os registradores de dados que treinam as redes neurais e os engenheiros que gerenciam a base de conhecimento.

Na saída, as redes neurais motoras transformam os dados categorizados em dados que controlam ações, como escrita de texto, síntese de imagens, saída de voz, instruções enviadas a quem responderá aos estímulos (robôs) etc. Esses neurônios motores são treinados com exemplos que correspondem aos dados categorizados da IEML com os dados motores. Novamente, os

dados de treinamento e as redes neurais devem ser diferenciados de acordo com seus tipos.

Memória e Processamento Semântico

A base de conhecimento é organizada por uma rede semântica; portanto, é preferencialmente apoiada por um *banco de dados gráfico*. Em termos de interface, essa base de conhecimento é apresentada como uma enciclopédia hipertextual. Também permite a programação de simulações e vários painéis de controle para monitoramento e inteligência.

O editor IEML mencionado na seção anterior também pode ser usado para tarefas que não sejam modelagem. Com efeito, permite operações variadas de leitura-escrita condicionadas pela presença de conteúdos semânticos localizados em determinadas funções gramaticais. Quando são codificados na IEML, os conceitos se tornam variáveis de uma álgebra, o que obviamente não é o caso quando são expressos em linguagem natural. Portanto, *transformações semânticas* podem ser programadas e computadas. Esta programação semântica abre o caminho não apenas para o *raciocínio lógico* clássico, ao qual os motores simbólicos de inferência da IA nos acostumaram há décadas, mas também a outras formas de raciocínio automático. Uma vez que na IEML a semântica é uma imagem funcional da sintaxe, torna-se possível automatizar *raciocínio analógico* como *A é para B o que C é para D*. Outras operações semânticas também podem ser programadas, tais como: seleção e pesquisa; substituição, inserção ou exclusão; extração de sub-redes semânticas relevantes; sumarização ou expansão; inversão, alusão, atenuação ou amplificação; extração ou projeção de estruturas narrativas etc.

Várias Aplicações

Algumas aplicações da nossa arquitetura de IA neuro-semântica IEML são evidentes: integração de dados; suporte de decisão com base em modelos causais; gestão do conhecimento; compreensão e síntese do texto; geração *controlada* de texto (ao contrário dos sistemas do tipo GPT3, nos quais a criação de texto não é controlada); chatbots e robótica. Agora comentaremos brevemente dois exemplos de aplicações: compreensão de texto e geração controlada de texto.

Em relação à *geração de texto controlado*, vamos imaginar dados de telemetria, informações contábeis, exames médicos, resultados de testes de conhecimento etc. como entrada (input). Como saída (output) podemos programar textos narrativos em linguagem natural sintetizando o conteúdo dos fluxos de dados de entrada: diagnósticos médicos, relatórios escolares, conselhos etc.

Sobre a *compreensão de texto*, vamos primeiro assumir a categorização automática do conteúdo de um documento inserido no sistema. Em uma segunda etapa, o modelo semântico extraído do texto é escrito na memória do sistema e integrado ao conhecimento que o sistema já adquiriu. Em suma, sistemas de Inteligência Artificial poderiam acumular conhecimento a partir da leitura automática de documentos. Supondo que a IEML seja adotada, os sistemas de Inteligência Artificial se tornariam capazes não apenas de acumular conhecimento, mas de integrá-lo em modelos coerentes e intercambiá-lo. Essa é obviamente uma perspectiva de longo prazo que exigirá esforços coordenados.

CONCLUSÃO: UM FUTURO HUMANÍSTICO PARA A IA

Mesmo que a arquitetura neuro-semântica proposta acima não contorne todos os obstáculos no caminho da Inteligência Artificial geral, ela favorecerá a IA na criação de aplicações capazes de processar o *significado* de textos ou situações. Ela também nos permite vislumbrar um mercado de dados rotulados em IEML que estimularia o já em expansão desenvolvimento do aprendizado de máquina. Também apoiaria uma memória pública colaborativa que seria particularmente útil nas áreas de pesquisa científica, educação e saúde.

Hoje, a multiplicidade de linguagens, sistemas de classificação, pontos de vista disciplinares e contextos práticos compartimenta nossa memória digital. No entanto, a comunicação de modelos, a comparação crítica dos pontos de vista e o acúmulo de conhecimento são essenciais para a cognição simbólica humana, uma cognição indissolivelmente pessoal e coletiva. A inteligência artificial só será capaz de aumentar de forma sustentável a cognição humana se for interoperável, cumulativa, integrável, intercambiável e distribuída. Isso significa que não faremos progressos significativos na Inteligência Artificial sem fazermos esforços simultâneos por uma inteligência coletiva capaz de autorreflexão e de se coordenar em uma memória global. A adoção de uma linguagem computável que funciona como um sistema universal de coordenadas semânticas – uma linguagem fácil de ler e escrever – abriria novos caminhos para a inteligência coletiva humana, incluindo uma interação multimídia imersiva no mundo das ideias. Nesse sentido, a comunidade de usuários da IEML pode ser o início de uma nova era de inteligência coletiva.

A IA contemporânea, a maioria estatística, tende a criar situações em que os dados pensam no lugar dos humanos, sem saber. Em contrapartida, ao adotar a IEML, propomos desenvolver uma IA que ajude os seres humanos a assumir o controle intelectual dos dados, a fim de extrair significado compartilhável, de forma sustentável. A IEML nos permite repensar o propósito e o funcionamento da IA de um ponto de vista humanista, um ponto de vista

para o qual o significado, a memória e a consciência pessoal devem ser tratados com a maior seriedade.

NOTAS E REFERÊNCIAS COMENTADAS

Sobre as Origens da IA

O termo *Inteligência Artificial* foi usado pela primeira vez em 1956 em uma conferência na Dartmouth College, em Hanover, New Hampshire. Os participantes da conferência incluíram o cientista da computação e pesquisador cognitivo Marvin Minsky (Turing Award 1969) e o inventor da linguagem de programação LISP, John McCarthy (Turing Award 1971).

Sobre o crescimento cognitivo

O *crescimento cognitivo* (em vez de *imitação* da inteligência humana) foi o foco principal de muitos pioneiros da ciência da computação e da Web. Ver, por exemplo:

- Bush, V. (1945, julho). As we may think. *Atlantic Monthly*.
- Licklider, J. (1960). Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics*, 1, 4-11.
- Engelbart, D. (1962). *Augmenting human intellect* [Relatório técnico]. Stanford Research Institute.
- Berners-Lee, T. (1999). *Weaving the Web*. Harper.

Sobre a História da IA Neural

Muitas pessoas reconhecem Geoffrey Hinton, Yann Le Cun e Yoshua Benjio como os fundadores da IA neural contemporânea. Mas a IA neural começou já na década de 1940. Uma breve bibliografia é fornecida a seguir.

- O primeiro artigo teórico sobre a IA neural foi publicado em 1943: McCulloch, W. S., & Pitts, W. (1943). A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115-133. <https://doi.org/10.1007/BF02478259>
- Warren McCulloch publicou vários artigos sobre esse tema que foram reunidos em: McCulloch, W. S. (1965). *Embodiments of mind*. MIT Press. Escrevi um artigo sobre seu trabalho: Lévy, P. (1986). L'Œuvre de Warren McCulloch. *Cahiers du CREA*, 7, 211-255.

- Frank Rosenblatt é o inventor do *Perceptron*, que pode ser considerado o primeiro sistema de aprendizagem de máquina baseado em uma rede neuromímica. Ver: Rosenblatt, F. (1962). *Principles of neurodynamics: Perceptrons and the theory of brain mechanisms*. Spartan Books.

- A tese de doutorado de Marvin Minsky, de 1954, foi intitulada *Theory of Neural-Analog Reinforcement Systems and Its Application to the Brain-Model Problem*. Minsky criticaria o *Perceptron* de Frank Rosenblatt em seu livro *Perceptrons* (MIT Press, 1969), escrito com Seymour Papert, e mais tarde continuaria o programa de pesquisa simbólico em IA. Também de Minsky, *The Society of Mind* (Simon and Schuster, 1986) resume bem sua abordagem da cognição humana como emergindo da interação de múltiplos módulos cognitivos com funções variadas.

- Heinz von Foerster foi secretário das Conferências Macy (1941-1960) sobre cibernética e teoria da informação. Foi diretor do Laboratório de Informática Biológica da University of Illinois (1958-1975). Seus principais artigos foram coletados em *Observing Systems: Selected Papers of Heinz von Foerster* (Intersystems Publications, 1981). Estudei de perto a pesquisa feita neste laboratório. Ver: Lévy, P. (1986). Analyse de contenu des travaux du Biological Computer Laboratory (BCL). *Cahiers du CREA*, 8, 155-191.

- Nos anos 1980, vamos notar a publicação do livro de referência de: McClelland, J. L., Rumelhart, D. E., & PDP Research Group. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. 2 vols. MIT Press.

- Desse mesmo ano, ver o importante artigo: Rumelhart, D. E., Hinton, G. E., Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533-536. <https://doi.org/10.1038/323533a0>

- Hinton foi mais tarde reconhecido por seu trabalho pioneiro com um Prêmio Turing junto com Yann LeCun e Joshua Benjio em 2018. Um de seus artigos mais citados é: LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436-444. <https://doi.org/10.1038/nature14539>

A Crítica da IA Estatística

Em relação à crítica da IA estatística, este texto retoma alguns dos argumentos apresentados por pesquisadores como Judea Pearl, Gary Marcus e Stephen Wolfram.

- Judea Pearl recebeu o Prêmio Turing em 2011 por seu trabalho sobre a modelagem da causalidade em IA. Ver: Pearl, J., & Mackenzie, D. (2019). *The book of why: The new science of cause and effect*. Basic Books.

- Gary Marcus escreveu em 2018 um artigo seminal: “Deep learning, a critical appraisal” (<https://arxiv.org/pdf/1801.00631.pdf?u>). Ver também: Marcus, G., & Davis, E. (2019). *Rebooting AI: Building artificial intelligence we can trust*. Vintage.
- Stephen Wolfram é o autor do software Mathematica e o cérebro por trás do mecanismo de busca Wolfram Alpha. Ver sua entrevista de 2016 para a Edge.org (https://www.edge.org/conversation/stephen_wolfram-ai-the-future-of-civilization).
- Além do trabalho de Judea Pearl sobre a importância da *modelagem causal* em IA, vamos lembrar as teses do filósofo Karl Popper sobre os limites do raciocínio indutivo e das estatísticas. Ver, em particular: Popper, K. (1972). *Objective knowledge: An evolutionary approach*. Clarendon Press.

Sobre a IA Neural Contemporânea

- Sobre o modelo de processamento de linguagem natural mais utilizado atualmente, chamado *BERT*, ver: [https://en.wikipedia.org/wiki/BERT_\(language_model\)](https://en.wikipedia.org/wiki/BERT_(language_model)).
- O recente relatório do Center for Research on Foundation Models (CRFM) do Stanford Institute for Human-Centered Artificial Intelligence (HAI), *On the Opportunities and Risks of Foundation Models*, começa com a frase: “A IA está passando por uma mudança de paradigma com o surgimento de modelos (por exemplo, BERT, DALL-E, GPT-3) que são treinados em dados amplos em escala e são adaptáveis a uma vasta gama de tarefas derivadas” (<https://arxiv.org/abs/2108.07258>).
- Sobre a Open IA Aberta, ver: <https://openai.com/blog/gpt-3-apps/> e <https://www.technologyreview.com/2020/08/22/1007539/gpt3-openai-language-generator-artificial-intelligence-ai-opinion/>.

Sobre IA Simbólica Contemporânea

- Integrar o conhecimento existente nos sistemas de IA é um dos principais objetivos da *Wolfram Language* de Stephen Wolfram. Ver <https://www.wolfram.com/language/principles/>.
- Sobre a Web Semântica, ver <https://www.w3.org/standards/semanticweb/#> e https://en.wikipedia.org/wiki/Semantic_Web.
- Sobre Wikidata ver: https://www.wikidata.org/wiki/Wikidata:Main_Page.
- A respeito do projeto Cyc de Douglas Lenat, ver: <https://en.wikipedia.org/wiki/Cyc>.

Sobre a Perspectiva Neuro-Simbólica

- Ernst, D. (2020). AI research and governance are at a crossroads. *CIGI Online*.
<https://www.cigionline.org/articles/ai-research-and-governance-are-crossroads/>

- Garcez, A. A., & Lamb, L. C. (2020). Neurosymbolic AI: The 3rd Wave.
<https://arxiv.org/pdf/2012.05876.pdf>

- Sobre a fusão neuro-simbólica, ver também o recente relatório da Universidade de Stanford *100 Year Study on AI*, que identifica a hipótese neuro-simbólica como uma das chaves para o avanço da disciplina (<https://ai100.stanford.edu/>).

Sobre interoperabilidade semântica

- Todos os editores de metadados semânticos afirmam ser interoperáveis, mas geralmente é uma interoperabilidade de *formatos de arquivo*, sendo estes últimos efetivamente assegurados pelos padrões da Web Semântica (XML, RDF, OWL etc.). Mas neste texto estou falando da interoperabilidade dos próprios *modelos semânticos* (de *arquiteturas de conceitos*: categorias e suas relações). É importante distinguir a interoperabilidade *semântica* da interoperabilidade do formato. Os modelos escritos na IEML podem ser exportados em formatos de metadados semânticos padrão, como RDF, JSON-LD ou Graph QL. Sobre a noção de interoperabilidade semântica, ver: <https://intlekt.io/2021/04/05/outline-of-a-business-model-for-a-change-in-civilization/>.

Sobre Chomsky e Sintaxe

Um dos primeiros pesquisadores a ter realizado uma matematização da linguagem é Noam Chomsky. Ver: Chomsky, N. (1957). *Syntactic structures*. Mouton e Chomsky, N., & Schützenberger, M.-P. (1963). The algebraic theory of context-free languages. In P. Braffort & D. Hirschberg (Eds.), *Computer Programming and Formal Languages* (pp. 118-161). Para uma abordagem mais filosófica, ver: Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge UP. Para entender como a IEML continua várias tendências da pesquisa linguística do século XX, ver: Lévy, P. (2021). The linguistic roots of IEML. *Intlekt*. <https://intlekt.io/the-linguistic-roots-of-ieml/>

Sobre Substantivos Próprios

Adoto aqui a posição de Saul Kripke, que é seguida pela maioria dos filósofos e gramáticos. Ver: Kripke, S. (1980). *Naming and Necessity*. Blackwell.

Ver minha recente entrada no blog sobre esse assunto: Lévy, P. (2021). Proper nouns in IEML. *Intlekt*. <https://intlekt.io/proper-names-in-ieml/>

Pierre Lévy sobre a IEML

- Lévy, P. (2009). Toward a self-referential collective intelligence: Some philosophical background of the IEML research program. In N. T. Nguyen, R. Kowalczyk & C. Shyi-Ming (Eds.), *Computational collective intelligence, semantic web, social networks and multiagent systems. First International Conference, ICCCI, Wroclaw, Poland, Oct. 2009, Proceedings* (pp. 22-35). Springer.

- Lévy, P. (2010, 2 de janeiro). The IEML research program: From social computing to reflexive collective intelligence. *Information sciences, special issue on collective intelligence*, 180(1), 71-94.

- As considerações filosóficas e científicas que me levaram à invenção do IEML foram amplamente descritas em: Lévy, P. (2011). *La sphere semantique. Computation, cognition, économie de l'information*. Hermes-Lavoisier. A tradução para o inglês: Lévy, P. (2011). *The semantic sphere. Computation, cognition and information economy*. Wiley. Esse livro contém uma extensa bibliografia.

- Os princípios gerais do IEML são resumidos em: <https://intlekt.io/ieml/>.

- Sobre a gramática IEML, ver: <https://intlekt.io/ieml-grammar/>.

- Sobre o dicionário IEML, ver: <https://intlekt.io/ieml-dictionary/>.

Outras referências relevantes de Pierre Lévy

- Lévy, P. (1994). *L'intelligence collective, pour une anthropologie du cyberspace*. La Découverte. Edição inglesa: Lévy, P. (1997). *Collective intelligence* (R. Bonono, Trans.). Perseus Books.

- Lévy, P. (1991). Les systèmes à base de connaissance comme médias de transmission de l'expertise. *Intellectica*, 12, 187-219.

- Analisei em detalhes o trabalho de engenharia do conhecimento em vários casos em meu livro: Lévy, P. (1992). *De la programmation considérée comme un des beaux-arts*. La Découverte. ■

Artigo recebido em 28 de março e aprovado em 15 de abril de 2022..